



COMPARAÇÃO DE DIFERENTES DESCRITORES DE IMAGENS DO CBERS 4A PARA A CLASSIFICAÇÃO DA COBERTURA DO SOLO COM UMA RANDOM FOREST

Jonilson Michel Fontes Galvão¹

Wanessa Monteiro Fernandes²

Antônio Cesar Germano Martins³

Darllan Collins da Cunha e Silva⁴

Recursos Hídricos e Qualidade da Água

Resumo

A análise e monitoramento da superfície terrestre por meio do sensoriamento remoto permite o estudo de grandes extensões de área. O presente trabalho tem o objetivo de comparar diferentes índices de vegetação como descritores a serem usados em uma *Random Forest* para a classificação de regiões em imagens do Satélite CBERS 4A, especificamente no entorno do Reservatório de Ituparanga, São Paulo. O processo de coleta de amostras ocorreu através do *Software Qgis* (com seleção de oito classes) e a implementação da *Random Forest* pelo *Rstudio*. Os resultados obtidos atestam que o NDVI possui maior contribuição na construção das árvores. A acurácia global registrada foi de 0,9433 e do Kappa de 0,9352. Este estudo contribui para o processo de tomada de decisão no âmbito da gestão ambiental por meio da automatização da classificação da cobertura do solo de forma eficiente.

Palavras-chave: Aprendizado supervisionado; Índices espectrais; Geoprocessamento.

¹ Doutorando em Ciências Ambientais – PPGCA, Universidade Estadual Paulista “Júlio de Mesquita Filho”, Programa de Pós-graduação em Ciências Ambientais, jonilson.galvao@unesp.br

² Mestra em Ciências Ambientais – PPGCA, Universidade Estadual Paulista “Júlio de Mesquita Filho”, Programa de Pós-graduação em Ciências Ambientais, wm.fernandes@unesp.br

³ Prof. Dr. Universidade Estadual Paulista “Júlio de Mesquita Filho” – Campus Sorocaba-SP, Instituto de Ciências e Tecnologia de Sorocaba - ICTS, antonio.martins@unesp.br

⁴ Prof. Dr. Universidade Estadual Paulista “Júlio de Mesquita Filho” – Campus Sorocaba-SP, Instituto de Ciências e Tecnologia de Sorocaba - ICTS, darllan.collins@unesp.br



INTRODUÇÃO

O estudo do uso do solo é fundamental para a gestão territorial e ambiental, tendo em vista a ocupação irregular de áreas, aliada a práticas irregulares e ao manejo incorreto, podendo ocasionar o empobrecimento do solo, a poluição e, muito além disso, o assoreamento de corpos hídricos (BERTONI & LOMBARDI NETO, 1999).

Através dos produtos do sensoriamento remoto é possível se calcular índices de vigor da vegetação que permitem supervisionar alterações naturais e antrópicas no uso e cobertura do solo. Esses índices podem ser calculados através de dados espectrais de duas ou mais bandas que são combinados entre si, com o intuito de associar os seus valores com indicadores biofísicos da vegetação (PONZONI, 2001).

Um índice muito utilizado é o que foi proposto por Rouse et al. (1973) o *Normalized Difference Vegetation Index* - NDVI (Índice de Vegetação por Diferença Normalizada). Calculado utilizando dados espectrais, esse índice mostra um bom desempenho para salientar mudanças na cobertura do solo, possibilitando que se avalie o teor de clorofila da vegetação através das bandas espectrais do infravermelho próximo e vermelho-visível.

Nesta mesma linha, outra possibilidade é a geração de índices utilizando canais de cores do espectro visível como feito por Ohta et al (1980) que analisaram índices baseados nos canais R (red), G (green) e B (blue), por meio das relações entre estes canais, constatando suas eficiências na segmentação de imagens coloridas.

Mas é no trabalho de Hamuda et al. (2016) que se encontra uma revisão de diferentes índices espectrais propostos na literatura e comparações dos resultados obtidos com esses índices quando aplicados à área agrícola.

Dessa forma, este trabalho tem como objetivo apresentar o resultado da utilização dos índices espectrais avaliados por Hamuda et al. (2016) como descritores a serem usados em uma *Random Forest* para apoio na gestão territorial, sendo a área de estudo, o entorno da represa Itupararanga, em São Paulo.

Realização



METODOLOGIA

A área de estudo compreende a represa de Itupararanga localizada na bacia do Alto Sorocaba criada em 1911 pela Companhia de Energia Elétrica – Light sendo alimentada pelos rios Sorocabaçu, Sorocamirim e Una (ROSA *et al.*, 2015). Além do abastecimento de água para diversas cidades, a represa contribui com a geração de energia elétrica e retém a água do Rio Sorocaba nos meses de cheia.

O reservatório tem capacidade de atender 800.000 habitantes, sendo responsável pelo fornecimento de água para os municípios de Sorocaba, Votorantim, Mairinque, Alumínio, Ibiúna e São Roque, todos pertencentes à região metropolitana de Sorocaba (HARKOT, 2019).

A região de estudo contém média de pluviosidade anual de 1493 mm, a classificação climática da área de estudo, com base na tabela de Koppen, é o subtropical úmido (Cwa) e possui médias de temperaturas acima de 22°, com estações seca e fria e quente e chuvosa. (SALLES *et al.* 2008).

Os índices avaliados por Hamuda *et al.* (2016) e aplicados neste estudo, são: NDI, ExG, ExR, CIVE, ExGR, NGRDI, VEG COM1, MExG, COM2, NDVI. A Figura 1 apresenta o esquema utilizado para o estudo da classificação por meio da *Random Forest* (RF).

A utilização desses índices como entrada de uma *Random Forest* (RF) foi realizada totalmente em ambiente de linguagem R, através da IDE Rstudio, na qual, por meio do pacote “randomForest” todo o procedimento é executado, sendo separado 70% das amostras para treino e 30% para validação. O estudo manteve as aplicações em modo padrão. Dessa forma, a taxa de erro *Out of Bag* (OOB) como procedimento imparcial, seleciona os melhores parâmetros da RF (BREIMAN, 2001; GOETZ *et al.* 2015; TRIGILA *et al.*, 2015).

Realização

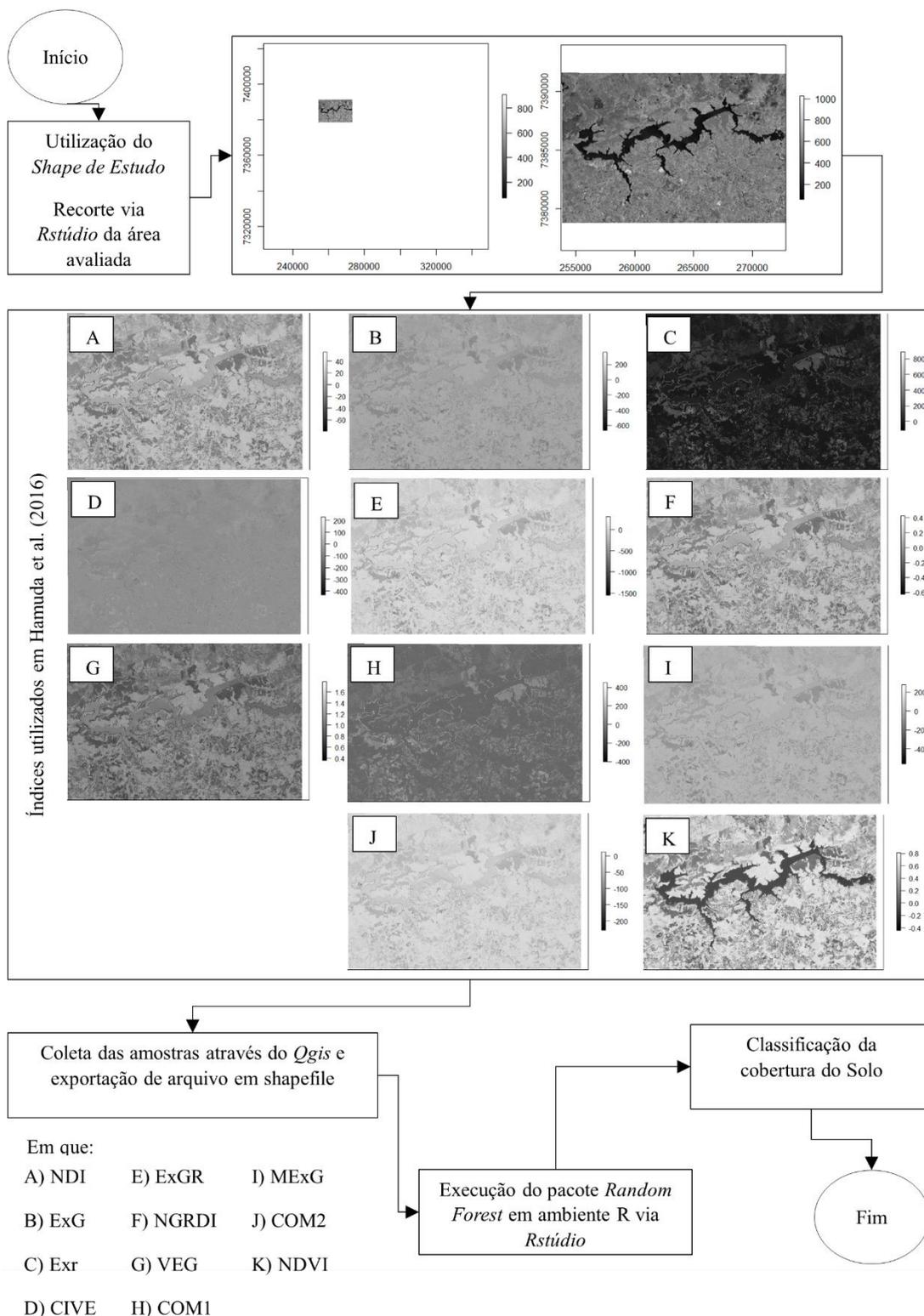


Figura 1. Esquema utilizado para o estudo da classificação por meio da RF.

Fonte: Elaborado pelos autores (2023)

Realização



RESULTADOS E DISCUSSÃO

A precisão geral da classificação e os coeficientes Kappa resultantes do *Stack* do conjunto das imagens CBERS 4A, utilizando as variáveis analisadas, são apresentados na Figura 2. Os resultados mostram que a classificação de cobertura do solo com base no conjunto de dados adquiridos em abril de 2023 obteve uma acurácia global de 0,9433 e um índice Kappa de 0,9352. Os maiores erros estão na em duas classes, Urbano e Solo exposto, que são similares aos encontrados nos estudos de Liu et al. (2019).

Predição	Referência								Total
	Agricultura	Água	Árborea	Eucalipto	Nuvem	Rasteira	Solo Exposto	Urbano	
Agricultura	703	0	12	0	0	8	0	31	754
Água	0	719	0	0	0	0	0	0	719
Árborea	23	0	723	8	3	8	0	17	782
Eucalipto	0	0	7	741	0	3	0	0	751
Nuvem	0	0	1	0	736	0	0	24	761
Rasteira	3	1	5	1	0	708	17	23	758
Solo Exposto	0	4	0	0	0	14	694	41	753
Urbano	21	3	2	0	11	9	39	614	699
Total	750	727	750	750	750	750	750	750	5977
	ACURÁCIA GLOBAL			0,9433	Kappa		0,9352		

Figura 2. Matriz de confusão para verificação dos métricas alcançadas no modelo RF

Fonte: Elaborado pelos autores (2023)

As importâncias dos descritores na geração das árvores são apresentados na Figura 3. Esse resultado é similar ao obtido por Hamuda et al. (2016), devido a maior contribuição no índice de vegetação NDVI. Na construção da RF, os índices mais relevantes foram o NDVI, CIVE, COM2 e COM1, ao passo que os coeficientes derivados da banda espectral verde apresentam menor contribuição.

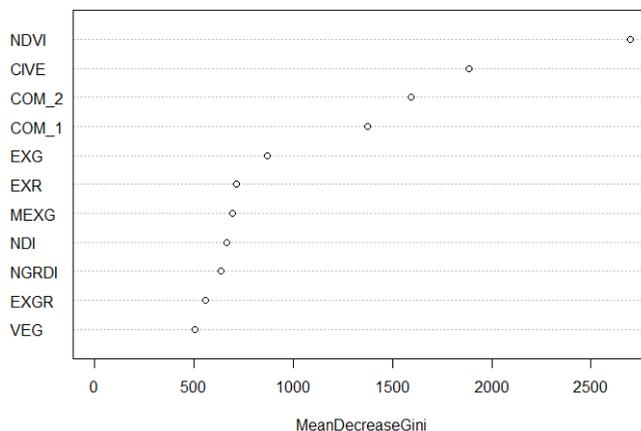


Figura 3. Variáveis de importância derivadas do algoritmo RF.
 Fonte: Elaborado pelos autores (2023)

As imagens do CBERS 4A classificada, CBERS 4A original e da região do Google Earth são apresentadas na Figura 4. As oito classes de cobertura do solo consideradas estão no mapa principal, com paletas únicas para cada uma. Observa-se maior confusão entre as classes de solo exposto e urbano na Figura 4.

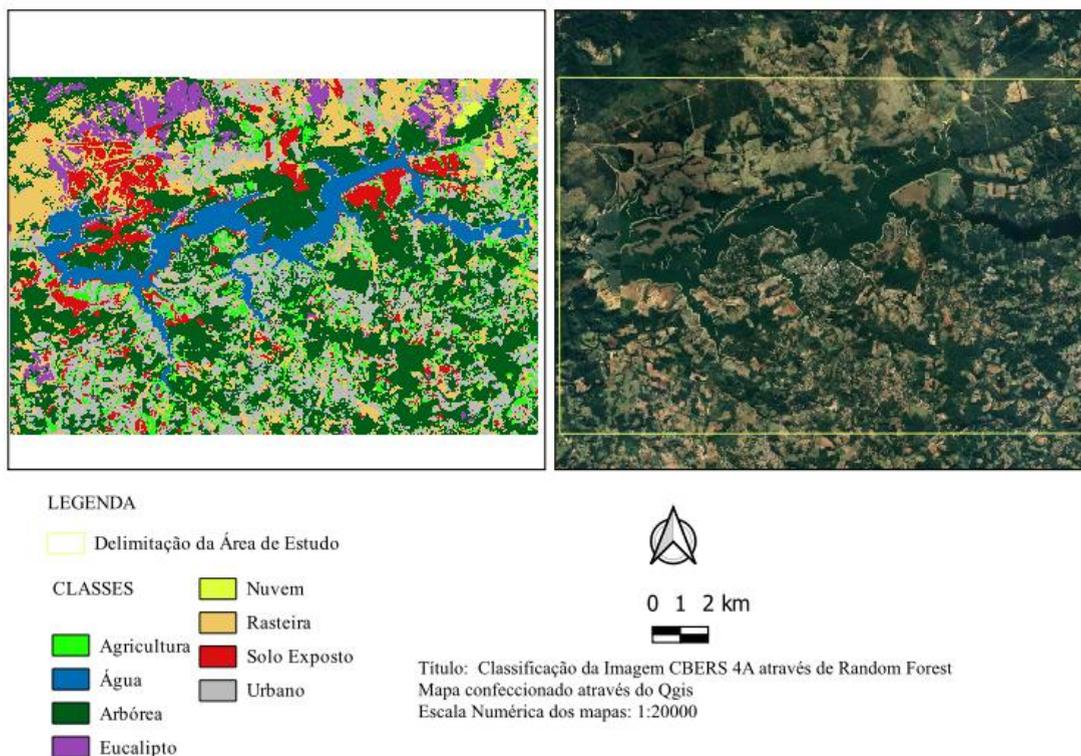


Figura 4. Mapa de classificação (à esquerda e acima) do conjunto de imagens CBERS 4A
 Fonte: Elaborado pelos autores (2023)

Realização



CONCLUSÕES

A utilização da aplicação do modelo Random Forest para a classificação das imagens do satélite CBERS 4A indica que o NDVI apresenta maior contribuição na construção das florestas aleatórias, ao passo que os índices menos relevantes são os derivados da banda verde do satélite.

Na classificação da área analisada, as classes que apresentam maior confusão para o modelo são o urbano e solo exposto. Ainda assim, a taxa geral de erro, torna viável a utilização do modelo para a tomada de decisão mais rápida e eficaz.

REFERÊNCIAS

HAMUDA, Esmael; GLAVIN, Martin; JONES, Edward. **A survey of image processing techniques for plant extraction and segmentation in the field.** Computers and Electronics in Agriculture, v. 125, p. 184-199, 2016.

HARKOT, O. A. M. G. **Conflitos socioambientais relacionados ao uso e ocupação do solo nas margens da represa Ituparanga – Ibúna, SP.** Dissertação (Mestrado Profissional em Administração, Gestão Ambiental e Sustentabilidade). Programa de Mestrado em Administração, Gestão Ambiental e Sustentabilidade, Universidade Nove de Julho. São Paulo, p. 76. 2019.

LIU, Qingsheng; SONG, Hongwei; LIU, Gaohuan; HUANG, Chong; LI, He. **Evaluating the Potential of Multi-Seasonal CBERS-04 Imagery for Mapping the Quasi-Circular Vegetation Patches in the Yellow River Delta Using Random Forest.** Remote Sensing, v. 11, n. 10, p. 1216, 22 maio 2019.

OHTA, Yu-Ichi; KANADE, Takeo; SAKAI, Toshiyuki. **Color information for region segmentation.** Computer graphics and image processing, v. 13, n. 3, p. 222-241, 1980.

PONZONI, F. J. **Comportamento espectral da vegetação.** In. Sensoriamento Remoto - Reflectância dos alvos naturais. Brasília: UnB, EMBRAPA, 2001. p.157-199. R. Nemani, C.D. Keeling, H. Hashimoto, W.M. Jolly, S.C. Piper, C.J. Tucker, R.B. Myneni, S.W. Running Climate-driven increases in global terrestrial net primary production from 1982 to 1999. Science, p. 1560-1563, v. 300, 2003.

ROSA, A. H.; SILVA, Â. A. M. J.; MELO, C. de A., CARLOS, V. M.; GUANDIQUE, M. E. G., FRACETO, L. F.; LOURENÇO, R. W. **Diagnóstico ambiental e avaliação de uso e ocupação do solo visando a sustentabilidade da represa de Ituparanga, importante área da bacia do médio Tietê.** In: POMPÊO, M.; MOSCHINI-CARLOS, V.; NISHIMURA, P. Y.; SILVA, S. C. da; DOVAL, J. C. L.. Ecologia de reservatórios e interfaces. São Paulo: Instituto de Biociências

Realização



da Universidade de São Paulo, ed. xii, p. 212-231, 2015.

ROUSE JR, J. W. et al. **MONITORING VEGETATION SYSTEMS IN THE GREAT PLAINS WITH ERTS**. In: Third Earth Resources Technology Satellite-1 Symposium: Section AB. Technical presentations. Scientific and Technical Information Office, National Aeronautics and Space Administration, p. 309, 1973.

Realização